1600

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/005,318E

DATE: 07/30/2002
TIME: 15:12:02

Input Set : A:\Epi3004b.app
Output Set: N:\CRF3\07302002\I005318E.raw

<110> APPLICANT: HEIN, MICH B.
HIATT, ANDREW C.
FITCHEN, JOHN H.

<120> TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT

<130> FILE REFERENCE: EPI3004B

<140> CURRENT APPLICATION NUMBER: 09/005,318E
<141> CURRENT FILING DATE: 1998-01-09
<150> PRIOR APPLICATION NUMBER: 08/782,481
<151> PRIOR FILING DATE: 1997-01-10
<150> PRIOR APPLICATION NUMBER: 09/005,167
<151> PRIOR FILING DATE: 1998-01-09

<160> NUMBER OF SEQ ID NOS: 113

<170> SOFTWARE: PatentIn Ver. 2.1

<210> SEQ ID NO: 1
<211> LENGTH: 137
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

```
Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys Ala
 1               5                  10                  15
Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu Asp
            20                  25                  30
Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg Glu
        35                  40                  45
Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Pro Val Tyr His
    50                  55                  60
Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu Asp
65                  70                  75                  80
Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp Ser
                85                  90                  95
Ala Thr Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Ala
            100                 105                 110
Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr Ala
        115                 120                 125
Leu Thr Pro Asp Ala Cys Tyr Pro Asp
    130                 135
```

<210> SEQ ID NO: 2
<211> LENGTH: 135
<212> TYPE: PRT
<213> ORGANISM: Mus sp.

<400> SEQUENCE: 2

```
Gln Asp Glu Asn Glu Arg Ile Val Val Asp Asn Lys Cys Lys Cys Ala
 1               5                  10                  15
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/005,318E

DATE: 07/30/2002
TIME: 15:12:02

Input Set : A:\Epi3004b.app
Output Set: N:\CRF3\07302002\I005318E.raw

```
67 Arg Ile Thr Ser Arg Ile Ile Pro Ser Ala Glu Asp Pro Ser Gln Asp
68              20                  25                  30
70 Ile Val Glu Arg Asn Val Arg Ile Ile Val Pro Leu Asn Ser Arg Glu
71          35                  40                  45
73 Asn Ile Ser Asp Pro Thr Ser Pro Met Arg Thr Lys Pro Val Tyr His
74      50                  55                  60
76 Leu Ser Asp Leu Cys Lys Lys Cys Asp Thr Thr Glu Val Glu Leu Glu
77  65                  70                  75                  80
79 Asp Gln Val Val Thr Ala Ser Gln Ser Asn Ile Cys Asp Ser Asp Ala
80                  85                  90                  95
82 Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Asn Arg Val
83              100                 105                 110
85 Lys Leu Ser Tyr Arg Gly Gln Thr Lys Met Val Glu Thr Ala Leu Thr
86          115                 120                 125
88 Pro Asp Ser Cys Tyr Pro Asp
89      130                 135
92 <210> SEQ ID NO: 3
93 <211> LENGTH: 137
94 <212> TYPE: PRT
95 <213> ORGANISM: Oryctolagus cuniculus
97 <400> SEQUENCE: 3
98 Asp Asp Glu Ala Thr Ile Leu Ala Asp Asn Lys Cys Met Cys Thr Arg
99   1               5                  10                  15
101 Val Thr Ser Arg Ile Ile Pro Ser Thr Glu Asp Pro Asn Glu Asp Ile
102             20                  25                  30
104 Val Glu Arg Asn Ile Arg Ile Val Val Pro Leu Asn Asn Arg Glu Asn
105         35                  40                  45
107 Ile Ser Asp Pro Thr Ser Pro Leu Arg Arg Asn Pro Val Tyr His Leu
108     50                  55                  60
110 Ser Asp Val Cys Lys Lys Cys Asp Pro Val Glu Val Glu Leu Glu Asp
111  65                  70                  75                  80
113 Gln Val Val Thr Ala Thr Gln Ser Asn Ile Cys Asn Glu Asp Asp Gly
114                 85                  90                  95
116 Val Pro Glu Thr Cys Tyr Met Tyr Asp Arg Asn Lys Cys Tyr Thr Thr
117             100                 105                 110
119 Met Val Pro Leu Arg Tyr His Gly Glu Thr Lys Met Val Gln Ala Ala
120         115                 120                 125
122 Leu Thr Pro Asp Ser Cys Tyr Pro Asp
123     130                 135
126 <210> SEQ ID NO: 4
127 <211> LENGTH: 136
128 <212> TYPE: PRT
129 <213> ORGANISM: Bos sp.
131 <400> SEQUENCE: 4
132 Glu Asp Glu Ser Thr Val Leu Val Asp Asn Lys Cys Gln Cys Val Arg
133   1               5                  10                  15
135 Ile Thr Ser Arg Ile Ile Arg Asp Pro Asp Asn Pro Ser Glu Asp Ile
136             20                  25                  30
138 Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Thr Arg Glu Asn
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/005,318E

DATE: 07/30/2002
TIME: 15:12:02

Input Set : A:\Epi3004b.app
Output Set: N:\CRF3\07302002\I005318E.raw

```
                35                  40                  45
Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Glu Pro Lys Tyr Asn Leu
     50                  55                  60
Ala Asn Leu Cys Lys Lys Cys Asp Pro Thr Glu Ile Glu Leu Asp Asn
 65                  70                  75                  80
Gln Val Phe Thr Ala Ser Gln Ser Asn Ile Cys Pro Asp Asp Asp Tyr
                85                  90                  95
Ser Glu Thr Cys Tyr Thr Tyr Asp Arg Asn Lys Cys Tyr Thr Thr Leu
            100                 105                 110
Val Pro Ile Thr His Arg Gly Val Thr Arg Met Val Lys Ala Thr Leu
        115                 120                 125
Thr Pro Asp Ser Cys Tyr Pro Asp
    130                 135
```

<210> SEQ ID NO: 5
<211> LENGTH: 119
<212> TYPE: PRT
<213> ORGANISM: Rana sp.
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (47)
<223> OTHER INFORMATION: Variable amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (88)..(89)
<223> OTHER INFORMATION: Variable amino acid
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (91)
<223> OTHER INFORMATION: Variable amino acid
<400> SEQUENCE: 5

```
Glu Gln Glu Tyr Ile Leu Ala Asn Asn Lys Cys Lys Cys Val Lys Ile
 1               5                  10                  15
Ser Ser Arg Phe Val Pro Ser Thr Glu Arg Pro Gly Glu Glu Ile Leu
            20                  25                  30
Glu Arg Asn Ile Gln Ile Thr Ile Pro Thr Ser Ser Arg Met Xaa Ile
        35                  40                  45
Ser Asp Pro Tyr Ser Pro Leu Arg Thr Gln Pro Val Tyr Asn Leu Trp
    50                  55                  60
Asp Ile Cys Gln Lys Cys Asp Pro Val Gln Leu Glu Ile Gly Gly Ile
65                  70                  75                  80
Pro Val Leu Ala Ser Gln Pro Xaa Xaa Ser Xaa Pro Asp Asp Glu Cys
                85                  90                  95
Tyr Thr Thr Glu Val Asn Phe Lys Lys Lys Val Pro Leu Thr Pro Asp
            100                 105                 110
Ser Cys Tyr Glu Tyr Ser Glu
        115
```

<210> SEQ ID NO: 6
<211> LENGTH: 128
<212> TYPE: PRT

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/005,318E

DATE: 07/30/2002
TIME: 15:12:02

Input Set : A:\Epi3004b.app
Output Set: N:\CRF3\07302002\I005318E.raw

209 <213> ORGANISM: Lumbricus sp.
211 <400> SEQUENCE: 6
212 Asn Lys Cys Met Cys Thr Arg Val Thr Ala Arg Ile Arg Gly Thr Arg
213 1 5 10 15
215 Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Tyr Ile Arg Ile Asn Val
216 20 25 30
218 Pro Leu Lys Asn Arg Gly Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg
219 35 40 45
221 Asn Gln Pro Val Tyr His Leu Ser Pro Ser Cys Lys Lys Cys Asp Pro
222 50 55 60
224 Tyr Glu Asp Gly Val Val Thr Ala Thr Glu Thr Asn Ile Cys Tyr Pro
225 65 70 75 80
227 Asp Gln Gly Val Pro Gln Ser Cys Arg Asp Tyr Cys Pro Glu Leu Asp
228 85 90 95
230 Arg Asn Lys Cys Tyr Thr Val Leu Val Pro Pro Gly Tyr Thr Gly Glu
231 100 105 110
233 Thr Lys Met Val Gln Asn Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
234 115 120 125
237 <210> SEQ ID NO: 7
238 <211> LENGTH: 421
239 <212> TYPE: DNA
240 <213> ORGANISM: Artificial Sequence
242 <220> FEATURE:
243 <221> NAME/KEY: CDS
244 <222> LOCATION: (1)..(414)
246 <220> FEATURE:
247 <221> NAME/KEY: sig_peptide
248 <222> LOCATION: (1)..(6)
250 <220> FEATURE:
251 <221> NAME/KEY: mat_peptide
252 <222> LOCATION: (7)..(414)
254 <220> FEATURE:
255 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
256 nucleotide sequence
258 <400> SEQUENCE: 7
259 gat cag gaa gat gaa cgt att gtt ctg gtt gac aac aag tgc aag tgt 48
260 Asp Gln Glu Asp Glu Arg Ile Val Leu Val Asp Asn Lys Cys Lys Cys
261 -1 1 5 10
263 gct cgt att act tct aga atc atc cgt agc tca gag gac cca aat gaa 96
264 Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser Ser Glu Asp Pro Asn Glu
265 15 20 25 30
267 gat ata gtc gaa cgt aac atc cgt atc atc gtc cca ctg aat aac cgg 144
268 Asp Ile Val Glu Arg Asn Ile Arg Ile Ile Val Pro Leu Asn Asn Arg
269 35 40 45
271 gag aat atc tca gat cct aca agt ccg ttg cgc aca cgc ttc gta tac 192
272 Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu Arg Thr Arg Phe Val Tyr
273 50 55 60
275 cac ctg tca gat ctg tgt aag aag tgt gat cca aca gag gta gag ctg 240
276 His Leu Ser Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/005,318E

DATE: 07/30/2002
TIME: 15:12:02

Input Set : A:\Epi3004b.app
Output Set: N:\CRF3\07302002\I005318E.raw

```
277          65                  70                  75
279 gac aat cag ata gtc act gcg act caa agc aac att tgc gat gag gac   288
280 Asp Asn Gln Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp
281      80                  85                  90
283 agc gct aca gaa acc tgc agc acc tac gat agg aac aaa tgc tac acg   336
284 Ser Ala Thr Glu Thr Cys Ser Thr Tyr Asp Arg Asn Lys Cys Tyr Thr
285  95                 100                 105                 110
287 gcc gtg gtt ccg ctc gtg tat ggt gga gag aca aaa atg gtg gaa act   384
288 Ala Val Val Pro Leu Val Tyr Gly Gly Glu Thr Lys Met Val Glu Thr
289                 115                 120                 125
291 gcc ctt acg ccc gat gca tgc tat ccg gac tgaattc                   421
292 Ala Leu Thr Pro Asp Ala Cys Tyr Pro Asp
293             130                 135
297 <210> SEQ ID NO: 8
298 <211> LENGTH: 215
299 <212> TYPE: DNA
300 <213> ORGANISM: Artificial Sequence
302 <220> FEATURE:
303 <221> NAME/KEY: CDS
304 <222> LOCATION: (1)..(213)
306 <220> FEATURE:
307 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
308       nucleotide sequence
310 <400> SEQUENCE: 8
311 gat cag aag tgc aag tgt gct cgt att act tct aga atc atc cgt agc   48
312 Asp Gln Lys Cys Lys Cys Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser
313   1               5                  10                  15
315 tca gag gac cca aat gaa gat ata gtc gaa cgt aac atc cgt atc atc   96
316 Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile
317             20                  25                  30
319 gtc cca ctg aat aac cgg gag aat atc tca gat cct aca agt ccg ttg   144
320 Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu
321          35                  40                  45
323 cgc aca cgc ttc gta tac cac ctg tca gat ctg tgt aag aag gat gag   192
324 Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Asp Glu
325      50                  55                  60
327 gac agc gct aca gaa acc tgc tg                                   215
328 Asp Ser Ala Thr Glu Thr Cys
329  65                  70
333 <210> SEQ ID NO: 9
334 <211> LENGTH: 140
335 <212> TYPE: DNA
336 <213> ORGANISM: Artificial Sequence
338 <220> FEATURE:
339 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
340       nucleotide sequence
342 <400> SEQUENCE: 9
343 ctagaatcat ccgtagctca gaggacccaa atgaagatat agtcgaacgt aacatccgta 60
344 tcatcgtccc actgaataac cgggagaata tctcagatcc tacaagtccg ttgcgcacac 120
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/30/2002
PATENT APPLICATION: US/09/005,318E TIME: 15:12:03

Input Set : A:\Epi3004b.app
Output Set: N:\CRF3\07302002\I005318E.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; Xaa Pos. 47,88,89,91

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/005,318E

DATE: 07/30/2002
TIME: 15:12:03

Input Set : A:\Epi3004b.app
Output Set: N:\CRF3\07302002\I005318E.raw

L:187 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:32
L:196 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:80